UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/036,729 | 12/21/2001 | Jaap M. Middeldorp | 9310-13DVCTDV | 6359 |

20792 7590 10/16/2007
MYERS BIGEL SIBLEY & SAJOVEC
PO BOX 37428
RALEIGH, NC 27627

| EXAMINER | |
|---|---|
| LI, QIAN JANICE | |
| **ART UNIT** | **PAPER NUMBER** |
| 1633 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 10/16/2007 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/036,729 | MIDDELDORP ET AL. | |
| | Examiner | Art Unit | |
| | Q. Janice Li, M.D. | 1633 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *23 July 2007*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *6-9,26,27 and 32-34* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *6-9,26,27 and 32-34* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

## DETAILED ACTION

The amendment and response filed 7/23/2007 are acknowledged. Claim 34 has been amended. Claims 6-9, 26, 27, 32-34 are pending and under current examination.

Unless otherwise indicated, previous rejections that have been rendered moot in view of the amendment to pending claims or persuasive arguments will not be reiterated. The arguments in 7/23/07 response would be addressed to the extent that they apply to current rejection.

### *Claim Rejections - 35 USC § 112*

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 6, 9 stand rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

In the remarks, the applicant first asserts,

> In claim 6, the nucleic acid sequences encode peptides comprising an epitope of the VCA-p18 or VCA-p40 protein.

In response, it is noted claim 6 as written, the nucleic acid sequence encodes a peptide that is immunochemically reactive with antibodies to VCA-18 or VCA-p40. Hence, the genus embraced by the claimed nucleic acids may encompass coding sequences for peptides comprising an epitope of the VCA-p18 or VCA-p40, but they are not limited by such coding sequence, rather the genus encompasses nucleic acid sequences encoding any protein as long as it is immunochemically reactive with antibodies to VCA-18 or VCA-p40, which include proteins having no clear relationship to VCA proteins. Claim 6 contains multiple phrases further limiting the antibodies, but not further limiting the peptide itself. The only limitation to the peptide is its immunochemical reactivity to the antibodies. One cannot extrapolate the structures of the protein genus from the immune reactivity.

Applicants then argues that the antibodies are further defined by particular hybridoma, and one of skill in the art further recognizes that monoclonal antibodies are specific to a single epitope and all of the antibodies produced from a single hybridoma are identical.

In response, it is noted that the antibodies are not limited by the recited hybridomas, but rather limited by the immunochemical reactivity: "*are antibodies having the same reactivity with VCA-p18 as antibodies produced by the hybridoma...*".

Even assuming *arguendo* claims are directed to peptides reactive with the antibodies produced by the recited hybridoma, the peptides encompass a genus of amino acid sequences, whose nucleic acid coding sequences are limited only by the fact that they contain a nucleotide, whose expression product has certain immune

reactivity, while since there is no length limitation on the nucleic acid sequences, they could be a genomic sequence or a fusion protein sequence, or any other known or unknown nucleic acid sequences that happen to encode a protein reactive to the antibodies or contain sequences sharing homology with SEQ ID Nos: 1 and/or 3. Thus, even if the antibodies produced by the hybridoma have the same reactivity with VCA-18 or VCA-40, the proteins reactive to the antibodies may be structurally distinct. This could be seen in the example of the specification, where serum proteins reactive with applicant's antibody have different domain combinations, i.e structurally different (see e.g. table I of the specification).

Applicants then argue that the specification demonstrates actual reduction to practice of the nucleic acids of claim 6 as exemplified by peptides represented by various SEQ ID Nos: 2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22.

In response, it is noted that majority of the numbered peptides are only a partial sequence of a protein, which each has an amino acid sequence distinct from the other, yet they all demonstrated immune reactivity with a VCA-p18 antibody. Among the 15 proteins listed in table I, only proteins from serum No. 9 and 10 have the same combination of domains I, II, & III, yet still it's unclear whether they have the same protein sequences in the remaining portion of the immune reactive protein. One skilled in the art could not extrapolate the full-length protein sequences from the disclosure of the partial sequences and their immune reactivity. Thus, working examples of the specification, especially table I further supports the Office position, the specification fails

to provide an adequate description for the genus of the nucleic acid sequences encompassed by the claims.

An adequate written description for a genus of nucleic acid sequences encompassed by instant claims requires more than a mere statement that they are part of the invention, what is required is a description of the sequences themselves. The court has made it very clear "CONCEPTION OF CHEMICAL COMPOUND REQUIRES THAT INVENTOR BE ABLE TO DEFINE COMPOUND SO AS TO DISTINGUISH IT FROM OTHER MATERIALS, AND TO DESCRIBE HOW TO OBTAIN IT, RATHER THAN SIMPLY DEFINING IT SOLELY BY ITS PRINCIPAL BIOLOGICAL ACTIVITY". *Amgen Inc. v. Chugai Pharmaceutical Co. Ltd.*, 18 USPQ2d 1016 (Fed. Cir. 1991).

The Revised Interim Guidelines state "THE CLAIMED INVENTION AS A WHOLE MAY NOT BE ADEQUATELY DESCRIBED IF THE CLAIMS REQUIRE AN ESSENTIAL OR CRITICAL ELEMENT WHICH IS NOT ADEQUATELY DESCRIBED IN THE SPECIFICATION AND WHICH IS NOT CONVENTIONAL IN THE ART" (Column 3, page 71434), "WHEN THERE IS SUBSTANTIAL VARIATION WITHIN THE GENUS, ONE MUST DESCRIBE A SUFFICIENT VARIETY OF SPECIES TO REFLECT THE VARIATION WITHIN THE GENUS", "IN AN UNPREDICTABLE ART, ADEQUATE WRITTEN DESCRIPTION OF A GENUS WHICH EMBRACES WIDELY VARIANT SPECIES CANNOT BE ACHIEVED BY DISCLOSING ONLY ONE SPECIES WITHIN THE GENUS" (Column 2, page 71436).

*Vas-Cath Inc. v. Mahurkar*, 19USPQ2d 1111, clearly states that "APPLICANT MUST CONVEY WITH REASONABLE CLARITY TO THOSE SKILLED IN THE ART THAT, AS OF THE FILING DATE SOUGHT, HE OR SHE WAS IN POSSESSION *OF THE INVENTION*. THE INVENTION IS, FOR PURPOSES OF THE 'WRITTEN DESCRIPTION' INQUIRY, *WHATEVER IS NOW CLAIMED*." (See page 1117.) The

specification does not "clearly allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed." (See *Vas-Cath* at page 1116).

One cannot describe what one has not conceived. See *Fiddes v. Baird*, 30 USPQ2d 1481, 1483. In *Fiddes*, claims directed to mammalian FGF's were found to be unpatentable due to lack of written description for that broad class. The specification provided only the bovine sequence.

In view of these considerations, a skilled artisan would not have viewed the teachings of the specification as sufficient to show that the applicant was in possession of the claimed invention commensurate to its scope because it does not provide adequate written description for the genus sequences encompassed by the claims. Therefore, for reasons of record and set forth *supra*, the rejection stands.

Claims 6, 9 stand rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention, for reasons of record and set forth *supra.*

***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 6-9, 26, 27, 32-34 stand rejected under 35 U.S.C. 102(b) as being anticipated by *Laux et al.* (The EMBO J 1988;7:769-74).

*Laux et al.* teaches a nucleic acid sequence comprising a subsequence of instant SEQ ID No: 1 (residues 1-535 of SEQ ID No: 1), which encodes at least 12 contiguous amino acids of EBV VCA-p18 (the amino acid sequence SEQ No: 5), which would be immunochemically reactive with antibodies to the EBV VCA-p18. *Laux et al.* also teaches a nucleic acid sequence comprising instant SEQ ID No: 3, which encodes 12 contiguous amino acids of an EBV VCA-40. *Laux et al.* discloses a vector comprising the sequences (e.g. fig. 1). Accordingly, *Laux et al.* anticipates instant claims.

In the remarks, the applicant asserts that multiple alignments were performed and homology as stated in the Office action was not found.

In response, the printed copy of the database search result is hereby enclosed, which contains the details of the alignment.

Claims 6-9, 26, 27, 32-34 stand rejected under 35 U.S.C. 102(b) as being anticipated by *Bankier et al.* (Mol Biol Med 1983;1:425-445).

*Bankier et al.* teaches a nucleic acid sequence comprising a subsequence of instant SEQ ID No: 1 (residues 1-535 of SEQ ID No: 1), which encodes at least 12 contiguous amino acids of EBV VCA-p18 (the amino acid sequence SEQ No: 5), which would be immunochemically reactive with antibodies to the EBV VCA-p18. *Bankier et al.* also teaches a nucleic acid sequence comprising a sequence that shares 98.8%

homology with instant SEQ ID No: 3 (subsequence thereof), which encodes 12 contiguous amino acids of an EBV VCA-40. *Bankier et al.* discloses that the sequences were cloned in a vector (e.g. figs. 1-3). Accordingly, *Bankier et al.* anticipates instant claims.

In the remarks, the applicant asserts that multiple alignments were performed and homology as stated in the Office action was not found.

In response, the printed copy of the database search result is hereby enclosed, which contains details of the alignment.

No claim is allowed.

**THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to **Q. Janice Li** whose telephone number is **571-272-0730**. The examiner can normally be reached on 9:30 am - 7 p.m., Monday through Friday, except every other Wednesday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, **Joseph Woitach** can be reached on **571-272-0739**. The **fax** numbers for the organization where this application or proceeding is assigned are **571-273-8300**.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to (571) 272-0547. For all other customer support, please call the USPTO Call Center (UCC) at **800-786-9199**.

**Q. JANICE LI, M.D.**
**PRIMARY EXAMINER**

Q. Janice Li, M.D.
Primary Examiner
Art Unit 1633

*QJL*
October 10, 2007

SEQ 3 of 10/036,729

```
Qy        661 CACTCGGGGCCTTACGGATTTCAGCCTCATCAAAGCTACGAAGTGCCCAGATACGTCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137078 CACTCGGGGCCTTACGGATTTCAGCCTCATCAAAGCTACGAAGTGCCCAGATACGTCCCT 137137

Qy        721 CATCCGCCCCCACCACCAACTTCTCACCAGGCAGCTCAGGCGCAGCCTCCACCCCCGGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137138 CATCCGCCCCCACCACCAACTTCTCACCAGGCAGCTCAGGCGCAGCCTCCACCCCCGGGC 137197

Qy        781 ACACAGGCCCCCGAAGCCCACTGTGTGGCCGAGTCCACGATCCCTGAGGCGGGAGCAGCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137198 ACACAGGCCCCCGAAGCCCACTGTGTGGCCGAGTCCACGATCCCTGAGGCGGGAGCAGCC 137257

Qy        841 GGGAACTCTGGACCCCGGGAGGACACCAACCCTCAGCAGCCCACCACCGAGGGCCACCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137258 GGGAACTCTGGACCCCGGGAGGACACCAACCCTCAGCAGCCCACCACCGAGGGCCACCAC 137317

Qy        901 CGCGGAAAGAAACTGGTGCAGGCCTCTGCGTCCGGAGTGGCTCAGTCTAAGGAGCCCACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137318 CGCGGAAAGAAACTGGTGCAGGCCTCTGCGTCCGGAGTGGCTCAGTCTAAGGAGCCCACC 137377

Qy        961 ACCCCCAAGGCCAAGTCTGTGTCAGCCCACCTCAAGTCCATCTTTTGCGAGGAATTGCTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137378 ACCCCCAAGGCCAAGTCTGTGTCAGCCCACCTCAAGTCCATCTTTTGCGAGGAATTGCTG 137437

Qy       1021 AATAAACGCGTGGCTTGA 1038
              ||||||||||||||||||
Db     137438 AATAAACGCGTGGCTTGA 137455
```

RESULT 7
EBV

```
LOCUS       EBV                 172281 bp    DNA     circular VRL 18-APR-2005
DEFINITION  Epstein-Barr virus (EBV) genome, strain B95-8.
ACCESSION   V01555 J02070 K01729 K01730 V01554 X00498 X00499 X00784
VERSION     V01555.1  GI:59074
KEYWORDS    DNA polymerase; EBNA; genome; ribonucleotide reductase; tandem
            repeat; terminal repeat.
SOURCE      Human herpesvirus 4 (Epstein-Barr virus)
  ORGANISM  Human herpesvirus 4
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Gammaherpesvirinae; Lymphocryptovirus.
REFERENCE   1
  AUTHORS   Arrand,J.R., Rymo,L., Walsh,J.E., Bjorck,E., Lindahl,T. and
            Griffin,B.E.
  TITLE     Molecular cloning of the complete Epstein-Barr virus genome as a
            set of overlapping restriction endonuclease fragments
  JOURNAL   Nucleic Acids Res. 9 (13), 2999-3014 (1981)
   PUBMED   6269068
REFERENCE   2
  AUTHORS   Kozak,M.
  TITLE     Possible role of flanking nucleotides in recognition of the AUG
            initiator codon by eukaryotic ribosomes
  JOURNAL   Nucleic Acids Res. 9 (20), 5233-5252 (1981)
   PUBMED   7301588
REFERENCE   3
  AUTHORS   Deininger,P.L., Bankier,A., Farrell,P., Baer,R. and Barrell,B.
  TITLE     Sequence analysis and in vitro transcription of portions of the
            Epstein-Barr virus genome
  JOURNAL   J. Cell. Biochem. 19 (3), 267-274 (1982)
   PUBMED   6296170
REFERENCE   4
```

AUTHORS Farrell,P.J., Bankier,A., Seguin,C., Deininger,P. and Barrell,B.G.
TITLE Latent and lytic cycle promoters of Epstein-Barr virus
JOURNAL EMBO J. 2 (8), 1331-1338 (1983)
PUBMED 10872327

REFERENCE 5
AUTHORS Farrell,P.J., Deininger,P.L., Bankier,A. and Barrell,B.
TITLE Homologous upstream sequences near Epstein-Barr virus promoters
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 80 (6), 1565-1569 (1983)
PUBMED 6300857

REFERENCE 6 (bases 142687 to 159853)
AUTHORS Bankier,A.T., Deininger,P.L., Farrell,P.J. and Barrell,B.G.
TITLE Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8 Epstein-Barr virus
JOURNAL Mol. Biol. Med. 1 (1), 21-45 (1983)
PUBMED 6092825

REFERENCE 7 (bases 112620 to 125316)
AUTHORS Seguin,C., Farrell,P.J. and Barrell,B.G.
TITLE DNA sequence and transcription of the BamHI fragment B region of B95-8 Epstein-Barr virus
JOURNAL Mol. Biol. Med. 1 (3), 369-392 (1983)
PUBMED 6094953

REFERENCE 8 (bases 45644 to 52450)
AUTHORS Jeang,K.T. and Hayward,S.D.
TITLE Organization of the Epstein-Barr virus DNA molecule. III. Location of the P3HR-1 deletion junction and characterization of the NotI repeat units that form part of the template for an abundant 12-O-tetradecanoylphorbol-13-acetate-induced mRNA transcript
JOURNAL J. Virol. 48 (1), 135-148 (1983)
PUBMED 6310141

REFERENCE 9 (bases 159853 to 172281)
AUTHORS Bankier,A.T., Deininger,P.L., Satchwell,S.C., Baer,R., Farrell,P.J. and Barrell,B.G.
TITLE DNA sequence analysis of the EcoRI Dhet fragment of B95-8 Epstein-Barr virus containing the terminal repeat sequences
JOURNAL Mol. Biol. Med. 1 (4), 425-445 (1983)
PUBMED 6094955

REFERENCE 10 (bases 45415 to 52824)
AUTHORS Jones,M.D., Foster,L., Sheedy,T. and Griffin,B.E.
TITLE The EB virus genome in Daudi Burkitt's lymphoma cells has a deletion similar to that observed in a non-transforming strain (P3HR-1) of the virus
JOURNAL EMBO J. 3 (4), 813-821 (1984)
PUBMED 6327290

REFERENCE 11 (bases 87650 to 92703)
AUTHORS Biggin,M., Farrell,P.J. and Barrell,B.G.
TITLE Transcription and DNA sequence of the BamHI L fragment of B95-8 Epstein-Barr virus
JOURNAL EMBO J. 3 (5), 1083-1090 (1984)
PUBMED 6203743

REFERENCE 12 (bases 7315 to 9312)
AUTHORS Yates,J., Warren,N., Reisman,D. and Sugden,B.
TITLE A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (12), 3806-3810 (1984)
PUBMED 6328526

REFERENCE 13 (bases 76089 to 79808)
AUTHORS Gibson,T., Stockwell,P., Ginsburg,M. and Barrell,B.
TITLE Homology between two EBV early genes and HSV ribonucleotide reductase and 38K genes
JOURNAL Nucleic Acids Res. 12 (12), 5087-5099 (1984)

PUBMED 6330697
REFERENCE 14 (bases 1 to 172281)
AUTHORS Baer,R., Bankier,A.T., Biggin,M.D., Deininger,P.L., Fallell,P.J., Gibson,T.J., Hatfull,G., Hudson,G.S., Satchwell,S.C., Seguin,C., Tuffnell,P.S. and Barrell,B.G.
TITLE DNA sequence and expression of the B95-8 Epstein-Barr virus genome
JOURNAL Nature 310 (5974), 207-211 (1984)
PUBMED 6087149
REFERENCE 15
AUTHORS Bodescot,M. and Perricaudet,M.
TITLE Clustered alternative splice sites in Epstein-Barr virus RNAs
JOURNAL Nucleic Acids Res. 15 (14), 5887 (1987)
PUBMED 3039467
REFERENCE 16
AUTHORS Laux,G., Perricaudet,M. and Farrell,P.J.
TITLE A spliced Epstein-Barr virus gene expressed in immortalized lymphocytes is created by circularization of the linear viral genome
JOURNAL EMBO J. 7 (3), 769-774 (1988)
PUBMED 2840285
REFERENCE 17
AUTHORS Hatfull,G.F., Barrell,B.G., Quinn,J. and McGeoch,D.
JOURNAL Unpublished
REFERENCE 18 (bases 1 to 172281)
AUTHORS Farrell,P.J. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-1984)
REFERENCE 19 (bases 1 to 172281)
AUTHORS Farrell,P.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1988) Farrell P., Ludwig Institute for Cancer Research, St. Mary's Hospital Medical School, Norfolk Place London W2 1PG
COMMENT On or before Apr 23, 2004 this sequence version replaced gi:330432, gi:330357, gi:330413.
CDS
Listed under this feature are all known protein coding regions as well as all the major open reading frames in the sequence. In general the term major is taken as the longest frame in a particular region taking into account the adjacent longest frames and likely transcription signals. Note that on this basis some long overlapping frames have been excluded and on the other hand some small frames have been included which might represent exons or genes because they occur in a logical combination with other features or because of some other experimental data. The reading frames are named according to the Bam H1 fragment in which they start. eg BALF3 is the third leftward frame starting in Bam H1 fragment A. BORF1 is the first rightward frame in Bam H1 fragment O. If there is an obvious TATA sequence followed by an in frame Met codon that satisfies the rules of Kozak [12] in that there is a purine at -3 and/or a G at +4 then the reading frame is numbered from the A of the ATG to the base preceding the termination codon. If there is no obvious initiation codon or there is a substantial reading frame in phase before the ATG then the reading frame is numbered from the first base of the first codon.
SITEs of POLYA signals
This feature lists all occurences of the sequence AATAAA which is found normally approximately 20 bases upstream of the mRNA processing/polyA addition site. The rarely used homolog ATTAAA is only listed when it is found in a position close to the end of a major reading frame.

```
            SITEs of DONOR and ACCEPT sequences
            This is not a comprehensive listing of all such sequences and only
            the positions of a few have been noted because they occur in
            potentially interesting positions. The number quoted in the table
            is the position of the terminal base in the intron in each case.
            Restriction enzyme SITEs.
            Only the positions of the sites Bam HI (BAM) are listed. RPT
            This feature is used to define repetitive sequences. SITE DEL
            This feature defines deletions in B95-8 with respect to other
            strains such as RAJI and also to deletions in other strains such as
            P3HR1 and DAUDI with respect to B95-8.
            SITE HPN
            Denotes sequences with twofold symmetry ie could form hairpin
            loops. This is not a comprehensive list - only a few occurences
            noted.
            ORGRPL
            Denotes the region that encompasses an origin of replication (ori
            P).[13].
            NUMBERING
            The DNA sequence of B95-8 EBV has been revised [19]. The original
            (Baer et al, 1984) base 359 has been deleted so the new sequence
            around that position reads TCAGTCTTT. To avoid renumbering the
            entire sequence, position 1 has benn moved 1 base to the left of
            the EcoRI site separating EcoRI Dhet from EcoRI I
            (ie the first A of AGAATTC).
FEATURES             Location/Qualifiers
     source          1. .172281
                     /organism="Human herpesvirus 4"
                     /mol_type="genomic DNA"
                     /strain="B95-8"
                     /db_xref="taxon:10376"
     mRNA            58. .272
                     /product="exon 2 terminal protein RNA"
     mRNA            360. .458
                     /product="exon 3 terminal protein RNA"
     misc_feature    complement(535)
                     /note="polyA signal: AATAAA"
     mRNA            540. .788
                     /product="exon 4 terminal protein RNA"
     mRNA            871. .951
                     /product="exon 5 terminal protein RNA"
     mRNA            1026. .1196
                     /product="exon 6 terminal protein RNA"
     promoter        complement(1192)
                     /note="TATA: TATAAAT"
     mRNA            1280. .1495
                     /product="exon 7 terminal protein RNA"
     promoter        complement(1383)

  Query Match             100.0%;  Score 1038;  DB 13;  Length 172281;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1038;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCTATCAGGTAACGCAGGAGAAGGAGCAACAGCCTGCGGAGGTTCGGCCGCCGCGGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148707 ATGCTATCAGGTAACGCAGGAGAAGGAGCAACAGCCTGCGGAGGTTCGGCCGCCGCGGGC 148766

Qy         61 CAGGACCTCATCAGCGTCCCCCGCAACACCTTTATGACACTGCTTCAGACCAACCTGGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148767 CAGGACCTCATCAGCGTCCCCCGCAACACCTTTATGACACTGCTTCAGACCAACCTGGAC 148826
```

```
Qy       121 AACAAACCGCCGAGGCAGACCCCGCTACCCTACGCGGCCCCGCTGCCCCCCTTTTCCCAC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148827 AACAAACCGCCGAGGCAGACCCCGCTACCCTACGCGGCCCCGCTGCCCCCCTTTTCCCAC 148886

Qy       181 CAGGCAATAGCCACCGCGCCTTCCTACGGTCCTGGGGCCGGAGCGGTCGCCCCGGCCGGC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148887 CAGGCAATAGCCACCGCGCCTTCCTACGGTCCTGGGGCCGGAGCGGTCGCCCCGGCCGGC 148946

Qy       241 GGCTACTTTACCTCCCCAGGAGGTTACTACGCCGGGCCCGCGGGCGGGGACCCGGGTGCC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148947 GGCTACTTTACCTCCCCAGGAGGTTACTACGCCGGGCCCGCGGGCGGGGACCCGGGTGCC 149006

Qy       301 TTCTTGGCGATGGACGCTCACACCTACCACCCCCACCCACACCCCCCTCCGGCCTACTTT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149007 TTCTTGGCGATGGACGCTCACACCTACCACCCCCACCCACACCCCCCTCCGGCCTACTTT 149066

Qy       361 GGCTTGCCGGGCCTCTTTGGCCCCCCTCCACCCGTGCCTCCTTACTACGGATCCCACTTG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149067 GGCTTGCCGGGCCTCTTTGGCCCCCCTCCACCCGTGCCTCCTTACTACGGATCCCACTTG 149126

Qy       421 CGGGCAGACTACGTCCCCGCTCCCTCGCGATCCAACAAGCGGAAAAGAGACCCCGAGGAG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149127 CGGGCAGACTACGTCCCCGCTCCCTCGCGATCCAACAAGCGGAAAAGAGACCCCGAGGAG 149186

Qy       481 GATGAAGAAGGCGGGGGGCTATTCCCGGGGGAGGACGCCACCCTCTACCGCAAGGACATA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149187 GATGAAGAAGGCGGGGGGCTATTCCCGGGGGAGGACGCCACCCTCTACCGCAAGGACATA 149246

Qy       541 GCGGGCCTCTCCAAGAGTGTGAATGAGTTACAGCACACGCTACAGGCCCTGCGCCGGGAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149247 GCGGGCCTCTCCAAGAGTGTGAATGAGTTACAGCACACGCTACAGGCCCTGCGCCGGGAG 149306

Qy       601 ACGCTGTCCTACGGCCACACCGGAGTCGGATACTGCCCCCAGCAGGGCCCCTGCTACACC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149307 ACGCTGTCCTACGGCCACACCGGAGTCGGATACTGCCCCCAGCAGGGCCCCTGCTACACC 149366

Qy       661 CACTCGGGGCCTTACGGATTTCAGCCTCATCAAAGCTACGAAGTGCCCAGATACGTCCCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149367 CACTCGGGGCCTTACGGATTTCAGCCTCATCAAAGCTACGAAGTGCCCAGATACGTCCCT 149426

Qy       721 CATCCGCCCCCACCACCAACTTCTCACCAGGCAGCTCAGGCGCAGCCTCCACCCCCGGGC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149427 CATCCGCCCCCACCACCAACTTCTCACCAGGCAGCTCAGGCGCAGCCTCCACCCCCGGGC 149486

Qy       781 ACACAGGCCCCCGAAGCCCACTGTGTGGCCGAGTCCACGATCCCTGAGGCGGGAGCAGCC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149487 ACACAGGCCCCCGAAGCCCACTGTGTGGCCGAGTCCACGATCCCTGAGGCGGGAGCAGCC 149546

Qy       841 GGGAACTCTGGACCCCGGGAGGACACCAACCCTCAGCAGCCCACCACCGAGGGCCACCAC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149547 GGGAACTCTGGACCCCGGGAGGACACCAACCCTCAGCAGCCCACCACCGAGGGCCACCAC 149606

Qy       901 CGCGGAAAGAAACTGGTGCAGGCCTCTGCGTCCGGAGTGGCTCAGTCTAAGGAGCCCACC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149607 CGCGGAAAGAAACTGGTGCAGGCCTCTGCGTCCGGAGTGGCTCAGTCTAAGGAGCCCACC 149666

Qy       961 ACCCCCAAGGCCAAGTCTGTGTCAGCCCACCTCAAGTCCATCTTTTGCGAGGAATTGCTG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149667 ACCCCCAAGGCCAAGTCTGTGTCAGCCCACCTCAAGTCCATCTTTTGCGAGGAATTGCTG 149726

Qy      1021 AATAAACGCGTGGCTTGA 1038
```

```
              ||||||||||||||||||
Db      149727 AATAAACGCGTGGCTTGA 149744


RESULT 8
HS4B958RAJ
LOCUS       HS4B958RAJ            184113 bp    DNA     linear   VRL 12-APR-1996
DEFINITION  Epstein-Barr virus, artifactual joining of B95-8 complete genome
            and the sequences from Raji of the large deletion found in B95-8.
ACCESSION   M80517 M75989
VERSION     M80517.1  GI:330330
KEYWORDS    .
SOURCE      Human herpesvirus 4 (Epstein-Barr virus)
  ORGANISM  Human herpesvirus 4
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Gammaherpesvirinae; Lymphocryptovirus.
REFERENCE   1  (sites)
  AUTHORS   Baer,R.J., Bankier,A.T., Biggin,M.D., Deininger,P.L., Farrell,P.J.,
            Gibson,T.J., Hatfull,G.F., Hudson,G.S., Satchwell,S.C., Seguin,C.,
            Tuffnell,P.S. and Barrell,B.G.
  TITLE     DNA sequence and expression of the B95-8 Epstein-Barr virus genome
  JOURNAL   Nature 310 (5974), 207-211 (1984)
   PUBMED   6087149
REFERENCE   2  (sites)
  AUTHORS   Parker,B.D., Bankier,A., Satchwell,S., Barrell,B. and Farrell,P.J.
  TITLE     Sequence and transcription of Raji Epstein-Barr virus DNA spanning
            the B95-8 deletion region
  JOURNAL   Virology 179 (1), 339-346 (1990)
   PUBMED   2171209
REFERENCE   3  (sites)
  AUTHORS   Sample,J., Brooks,L., Sample,C., Young,L., Rowe,M., Gregory,C.,
            Rickinson,A. and Kieff,E.
  TITLE     Restricted Epstein-Barr virus protein expression in Burkitt
            lymphoma is due to a different Epstein-Barr nuclear antigen 1
            transcriptional initiation site
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6343-6347 (1991)
   PUBMED   1648738
REFERENCE   4  (bases 1 to 184113)
  AUTHORS   Jenson,H.B.
  TITLE     GenBank Curator Program
  JOURNAL   Unpublished (1992)
COMMENT     Original source text: Human herpesvirus 4 DNA.
            The B95-8 genome (V01555) has a large deletion in the right side of
            the genome which has been sequenced in Raji (M35547).  These
            sequences have been joined to form an extended and more complete,
            although artifactual, EBV sequence.
            For features, refer to feature tables of V01555 and M35547.
FEATURES             Location/Qualifiers
     source          1. .184113
                     /organism="Human herpesvirus 4"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10376"
     misc_feature    1. .152008
                     /note="B95-8 sequences (corresponds to 1-152,008 of
                     V01555)"
     misc_feature    152009. .152012
                     /note="Overlap of B95-8 and Raji sequences at B95-8
                     deletion point (corresponds to 152,009-152,012 in V01555,
                     and 1-4 in M35547)"
     misc_feature    153013. .163839
                     /note="Raji sequences (corresponds to 5-11,831 of M35547)"
```

SEQ # 5

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=84270667; PubMed=6087149; DOI=10.1038/310207a0;
RA   Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA   Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA   Tuffnell P.S., Barrell B.G.;
RT   "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL   Nature 310:207-211(1984).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX   MEDLINE=86045997; PubMed=2998075;
RA   Hudson G.S., Gibson T.J., Barrell B.G.;
RT   "The BamHI F region of the B95-8 Epstein-Barr virus genome.";
RL   Virology 147:99-109(1985).
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; V01555; CAA24838.1; ALT_INIT; Genomic_DNA.
DR   EMBL; M11923; AAA45870.1; -; Genomic_DNA.
DR   InterPro; IPR009299; Herpes_capsid.
DR   Pfam; PF06112; Herpes_capsid; 1.
KW   Capsid protein; Virion protein.
FT   CHAIN         1    176       Capsid protein VP26.
FT                                /FTId=PRO_0000115739.
SQ   SEQUENCE   176 AA;  18147 MW;  DAB605ED00F1A656 CRC64;

  Query Match           100.0%;  Score 130;  DB 1;  Length 176;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-09;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVDTGSGGGGQPHDTAPRGARKKQ 24
              ||||||||||||||||||||||||
Db        153 AVDTGSGGGGQPHDTAPRGARKKQ 176


RESULT 2
Q777G5_EBVG
ID   Q777G5_EBVG    PRELIMINARY;   PRT;   176 AA.
AC   Q777G5;
DT   05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT   05-JUL-2004, sequence version 1.
DT   13-JUN-2006, entry version 11.
DE   Capsid protein VP26.
GN   Name=BFRF3;
OS   Epstein-Barr virus (strain GD1) (HHV-4) (Human herpesvirus 4).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10376;
OH   NCBI_TaxID=9606; Homo sapiens (Human).
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=84270667; PubMed=6087149; DOI=10.1038/310207a0;
RA   Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA   Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA   Tuffnell P.S., Barrell B.G.;
RT   "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL   Nature 310:207-211(1984).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
```

```
RC   STRAIN=B95-8;
RX   MEDLINE=88283646; PubMed=2840285;
RA   Laux G., Perricaudet M., Farrell P.J.;
RT   "A spliced Epstein-Barr virus gene expressed in immortalized
RT   lymphocytes is created by circularization of the linear viral
RT   genome.";
RL   EMBO J. 7:769-774(1988).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=82014887; PubMed=6269068;
RA   Arrand J.R., Rymo L., Walsh J.E., Bjorck E., Lindahl T., Griffin B.E.;
RT   "Molecular cloning of the complete Epstein-Barr virus genome as a set
RT   of overlapping restriction endonuclease fragments.";
RL   Nucleic Acids Res. 9:2999-3014(1981).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=82059504; PubMed=7301588;
RA   Kozak M.;
RT   "Possible role of flanking nucleotides in recognition of the AUG
RT   initiator codon by eukaryotic ribosomes.";
RL   Nucleic Acids Res. 9:5233-5252(1981).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=83109311; PubMed=6296170;
RA   Deininger P.L., Bankier A., Farrell P., Baer R., Barrell B.;
RT   "Sequence analysis and in vitro transcription of portions of the
RT   Epstein-Barr virus genome.";
RL   J. Cell. Biochem. 19:267-274(1982).
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=83169725; PubMed=6300857;
RA   Farrell P.J., Deininger P.L., Bankier A., Barrell B.;
RT   "Homologous upstream sequences near Epstein-Barr virus promoters.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:1565-1569(1983).
RN   [7]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=85035713; PubMed=6092825;
RA   Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;
RT   "Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8
RT   Epstein-Barr virus.";
RL   Mol. Biol. Med. 1:21-45(1983).
RN   [8]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=85060424; PubMed=6094953;
RA   Seguin C., Farrell P.J., Barrell B.G.;
RT   "DNA sequence and transcription of the BamHI fragment B region of B95-
RT   8 Epstein-Barr virus.";
RL   Mol. Biol. Med. 1:369-392(1983).
RN   [9]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=83294686; PubMed=6310141;
RA   Jeang K.T., Hayward S.D.;
RT   "Organization of the Epstein-Barr virus DNA molecule. III. Location of
RT   the P3HR-1 deletion junction and characterization of the NotI repeat
```

```
RT   units that form part of the template for an abundant 12-O-
RT   tetradecanoylphorbol-13-acetate-induced mRNA transcript.";
RL   J. Virol. 48:135-148(1983).
RN   [10]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=85060428; PubMed=6094955;
RA   Bankier A.T., Deininger P.L., Satchwell S.C., Baer R., Farrell P.J.,
RA   Barrell B.G.;
RT   "DNA sequence analysis of the EcoRI Dhet fragment of B95-8 Epstein-
RT   Barr virus containing the terminal repeat sequences.";
RL   Mol. Biol. Med. 1:425-445(1983).
RN   [11]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=20331131; PubMed=10872327;
RA   Farrell P.J., Bankier A., Seguin C., Deininger P., Barrell B.G.;
RT   "Latent and lytic cycle promoters of Epstein-Barr virus.";
RL   EMBO J. 2:1331-1338(1983).
RN   [12]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=84207939; PubMed=6327290;
RA   Jones M.D., Foster L., Sheedy T., Griffin B.E.;
RT   "The EB virus genome in Daudi Burkitt's lymphoma cells has a deletion
RT   similar to that observed in a non-transforming strain (P3HR-1) of the
RT   virus.";
RL   EMBO J. 3:813-821(1984).
RN   [13]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=84236104; PubMed=6203743;
RA   Biggin M., Farrell P.J., Barrell B.G.;
RT   "Transcription and DNA sequence of the BamHI L fragment of B95-8
RT   Epstein-Barr virus.";
RL   EMBO J. 3:1083-1090(1984).
RN   [14]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=84222045; PubMed=6328526;
RA   Yates J., Warren N., Reisman D., Sugden B.;
RT   "A cis-acting element from the Epstein-Barr viral genome that permits
RT   stable replication of recombinant plasmids in latently infected
RT   cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:3806-3810(1984).
RN   [15]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=84247360; PubMed=6330697;
RA   Gibson T.J., Stockwell P., Ginsburg M., Barrell B.G.;
RT   "Homology between two EBV early genes and HSV ribonucleotide reductase
RT   and 38K genes.";
RL   Nucleic Acids Res. 12:5087-5099(1984).
RN   [16]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=87289053; PubMed=3039467;
RA   Bodescot M., Perricaudet M.;
RT   "Clustered alternative splice sites in Epstein-Barr virus RNAs.";
RL   Nucleic Acids Res. 15:5887-5887(1987).
RN   [17]
```

```
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RX   MEDLINE=91021036; PubMed=2171209;
RA   Parker B.D., Bankier A., Satchwell S., Barrell B., Farrell P.J.;
RT   "Sequence and transcription of Raji Epstein-Barr virus DNA spanning
RT   the B95-8 deletion region.";
RL   Virology 179:339-346(1990).
RN   [18]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RA   Hatfull G.F., Barrell B.G., Quinn J., McGeoch D.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [19]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B95-8;
RA   Binne U.K., Amon W., Farrell P.J.;
RT   "Induction of Epstein-Barr virus late promoters on small plasmids in
RT   the EBV late lytic cycle requires ori lyt.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; AJ507799; CAD53401.1; -; Genomic_DNA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   InterPro; IPR009299; Herpes_capsid.
DR   Pfam; PF06112; Herpes_capsid; 1.
SQ   SEQUENCE   176 AA;  18147 MW;  DAB605ED00F1A656 CRC64;

  Query Match             100.0%;  Score 130;  DB 2;  Length 176;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-09;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AVDTGSGGGGQPHDTAPRGARKKQ 24
              ||||||||||||||||||||||||
Db        153 AVDTGSGGGGQPHDTAPRGARKKQ 176


RESULT 3
Q3KSU9_EBVG
ID   Q3KSU9_EBVG    PRELIMINARY;   PRT;   176 AA.
AC   Q3KSU9;
DT   08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT   08-NOV-2005, sequence version 1.
DT   11-JUL-2006, entry version 10.
DE   Hypothetical protein (BFRF3).
OS   Epstein-Barr virus (strain GD1) (HHV-4) (Human herpesvirus 4).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10376;
OH   NCBI_TaxID=9606; Homo sapiens (Human).
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=GD1;
RX   PubMed=16306603; DOI=10.1128/JVI.79.24.15323-15330.2005;
RA   Zeng M.-S., Li D.-J., Liu Q.-L., Song L.-B., Li M.-Z., Zhang R.-H.,
RA   Yu X.-J., Wang H.-M., Ernberg I., Zeng Y.-X.;
RT   "Genomic sequence analysis of Epstein-Barr virus strain GD1 from a
RT   nasopharyngeal carcinoma patient.";
RL   J. Virol. 79:15323-15330(2005).
RN   [2]
```

SEQ ID 07 10/036,729

```
Qy       481 TGGCGGGGGACAACCCCACGACACCGCCCCACGCGGGGCACGTAAGAAACAGTAG 535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7130 TGGCGGGGGACAACCCCACGACACCGCCCCACGCGGGGCACGTAAGAAACAGTAG 7184
```

RESULT 7
HHV507799
LOCUS HHV507799 171823 bp DNA circular VRL 15-APR-2005
DEFINITION Human herpesvirus 4 complete wild type genome.
ACCESSION AJ507799
VERSION AJ507799.1 GI:23893576
KEYWORDS complete genome.
SOURCE Human herpesvirus 4 (Epstein-Barr virus)
ORGANISM Human herpesvirus 4
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae; Lymphocryptovirus.
REFERENCE 1
AUTHORS Arrand,J.R., Rymo,L., Walsh,J.E., Bjorck,E., Lindahl,T. and Griffin,B.E.
TITLE Molecular cloning of the complete Epstein-Barr virus genome as a set of overlapping restriction endonuclease fragments
JOURNAL Nucleic Acids Res. 9 (13), 2999-3014 (1981)
PUBMED 6269068
REFERENCE 2
AUTHORS Kozak,M.
TITLE Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes
JOURNAL Nucleic Acids Res. 9 (20), 5233-5252 (1981)
PUBMED 7301588
REFERENCE 3
AUTHORS Deininger,P.L., Bankier,A., Farrell,P., Baer,R. and Barrell,B.
TITLE Sequence analysis and in vitro transcription of portions of the Epstein-Barr virus genome
JOURNAL J. Cell. Biochem. 19 (3), 267-274 (1982)
PUBMED 6296170
REFERENCE 4
AUTHORS Farrell,P.J., Bankier,A., Seguin,C., Deininger,P. and Barrell,B.G.
TITLE Latent and lytic cycle promoters of Epstein-Barr virus
JOURNAL EMBO J. 2 (8), 1331-1338 (1983)
PUBMED 10872327
REFERENCE 5
AUTHORS Farrell,P.J., Deininger,P.L., Bankier,A. and Barrell,B.
TITLE Homologous upstream sequences near Epstein-Barr virus promoters
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 80 (6), 1565-1569 (1983)
PUBMED 6300857
REFERENCE 6
AUTHORS Bankier,A.T., Deininger,P.L., Farrell,P.J. and Barrell,B.G.
TITLE Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8 Epstein-Barr virus
JOURNAL Mol. Biol. Med. 1 (1), 21-45 (1983)
PUBMED 6092825
REFERENCE 7
AUTHORS Seguin,C., Farrell,P.J. and Barrell,B.G.
TITLE DNA sequence and transcription of the BamHI fragment B region of B95-8 Epstein-Barr virus
JOURNAL Mol. Biol. Med. 1 (3), 369-392 (1983)
PUBMED 6094953
REFERENCE 8
AUTHORS Jeang,K.T. and Hayward,S.D.
TITLE Organization of the Epstein-Barr virus DNA molecule. III. Location

of the P3HR-1 deletion junction and characterization of the NotI repeat units that form part of the template for an abundant 12-O-tetradecanoylphorbol-13-acetate-induced mRNA transcript
JOURNAL J. Virol. 48 (1), 135-148 (1983)
PUBMED 6310141
REFERENCE 9
AUTHORS Bankier,A.T., Deininger,P.L., Satchwell,S.C., Baer,R., Farrell,P.J. and Barrell,B.G.
TITLE DNA sequence analysis of the EcoRI Dhet fragment of B95-8 Epstein-Barr virus containing the terminal repeat sequences
JOURNAL Mol. Biol. Med. 1 (4), 425-445 (1983)
PUBMED 6094955
REFERENCE 10
AUTHORS Jones,M.D., Foster,L., Sheedy,T. and Griffin,B.E.
TITLE The EB virus genome in Daudi Burkitt's lymphoma cells has a deletion similar to that observed in a non-transforming strain (P3HR-1) of the virus
JOURNAL EMBO J. 3 (4), 813-821 (1984)
PUBMED 6327290
REFERENCE 11
AUTHORS Biggin,M., Farrell,P.J. and Barrell,B.G.
TITLE Transcription and DNA sequence of the BamHI L fragment of B95-8 Epstein-Barr virus
JOURNAL EMBO J. 3 (5), 1083-1090 (1984)
PUBMED 6203743
REFERENCE 12
AUTHORS Yates,J., Warren,N., Reisman,D. and Sugden,B.
TITLE A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (12), 3806-3810 (1984)
PUBMED 6328526
REFERENCE 13
AUTHORS Gibson,T., Stockwell,P., Ginsburg,M. and Barrell,B.
TITLE Homology between two EBV early genes and HSV ribonucleotide reductase and 38K genes
JOURNAL Nucleic Acids Res. 12 (12), 5087-5099 (1984)
PUBMED 6330697
REFERENCE 14 (bases 1 to 171823)
AUTHORS Baer,R.J., Bankier,A.T., Biggin,M.D., Deininger,P.L., Farrell,P.J., Gibson,T.J., Hatfull,G.F., Hudson,G.S., Satchwell,S.C., Seguin,C., Tuffnell,P.S. and Barrell,B.G.
TITLE DNA sequence and expression of the B95-8 Epstein-Barr virus genome
JOURNAL Nature 310 (5974), 207-211 (1984)
PUBMED 6087149
REFERENCE 15
AUTHORS Bodescot,M. and Perricaudet,M.
TITLE Clustered alternative splice sites in Epstein-Barr virus RNAs
JOURNAL Nucleic Acids Res. 15 (14), 5887 (1987)
PUBMED 3039467
REFERENCE 16
AUTHORS Laux,G., Perricaudet,M. and Farrell,P.J.
TITLE A spliced Epstein-Barr virus gene expressed in immortalized lymphocytes is created by circularization of the linear viral genome
JOURNAL EMBO J. 7 (3), 769-774 (1988)
PUBMED 2840285
REFERENCE 17
AUTHORS Parker,B.D., Bankier,A., Satchwell,S., Barrell,B. and Farrell,P.J.
TITLE Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region

```
  JOURNAL   Virology 179 (1), 339-346 (1990)
   PUBMED   2171209
REFERENCE   18 (bases 1 to 171823)
  AUTHORS   Hatfull,G.F., Barrell,B.G., Quinn,J. and McGeoch,D.
  JOURNAL   Unpublished
REFERENCE   19
  AUTHORS   Binne,U.K., Amon,W. and Farrell,P.J.
  TITLE     Induction of Epstein-Barr virus late promoters on small plasmids in
            the EBV late lytic cycle requires ori lyt
  JOURNAL   Unpublished
REFERENCE   20 (bases 1 to 171823)
  AUTHORS   Farrell,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2002) Farrell P., Ludwig Institute for Cancer
            Research, Imperial College School of Medicine, St. Mary's Campus,
            Norfolk Place London W2 1PG
COMMENT     Construction:
            This sequence was assembled from B95-8 EBV [14] and Raji EBV [18]
            with sequence corrections [16, 19]. The number of major internal
            repeat units has been reduced from 11.6 [14] to a more typical 7.6
            and the B95-8 deletion sequences have been restored to give a
            sequence more representative of wild type EBV.
            Numbering
            Like the modified B95-8 sequence[14, 16] accession number V01555,
            this sequence starts 1 base to the left of the EcoRI site
            separating EcoRI Dhet from EcoRI I (ie the first A of AGAATTC.).
FEATURES             Location/Qualifiers
     source          1. .171823
                     /organism="Human herpesvirus 4"
                     /mol_type="genomic DNA"
                     /strain="B95-8"
                     /db_xref="taxon:10376"
     source          139724. .151554
                     /organism="Human herpesvirus 4"
                     /mol_type="genomic DNA"
                     /strain="Raji"
                     /db_xref="taxon:10376"
     CDS             join(166103. .166458,58. .272,360. .458,540. .788,871. .951,
                     1026. .1196,1280. .1495,1574. .1680)
                     /codon_start=1
                     /product="terminal protein LMP2A"
                     /protein_id="CAD53382.1"
                     /db_xref="GI:23893577"
                     /db_xref="InterPro:IPR010881"
                     /db_xref="UniProt/TrEMBL:Q777H4"
                     /translation="MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTP
                     PNDEERESNEEPPPPYEDPYWGNGDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYS
                     PRDDSSQHIYEEAGRGSMNPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALS
                     LLLLAAVASSYAAAQRKLLTPVTVLTAVVTFFAICLTWRIEDPPFNSLLFALLAAAGG
                     LQGIYVLVMLVLLILAYRRRWRRLTVCGGIMFLACVLVLIVDAVLQLSPLLGAVTVVS
                     MTLLLLAFVLWLSSPGGLGTLGAALLTLAAALALLASLILGTLNLTTMFLLMLLWTLV
                     VLLICSSCSSCPLSKILLARLFLYALALLLLASALIAGGSILQTNFKSLSSTEFIPNL
                     FCMLLLIVAGILFILAILTEWGSGNRTYGPVFMCLGGLLTMVAGAVWLTVMSNTLLSA
                     WILTAGFLIFLIGFALFGVIRCCRYCCYYCLTLESEERPPTPYRNTV"
     gene            join(58. .272,360. .458,540. .788,871. .951,1026. .1196,
                     1280. .1495,1574. .1682,5408. .5856,166103. .166458)
                     /gene="LMP2"
     exon            58. .272
                     /gene="LMP2"
                     /number=2
     CDS             join(59. .272,360. .458,540. .788,871. .951,1026. .1196,
```

```
                     1280. .1495,1574. .1680)
                     /gene="LMP2"
                     /codon_start=1
                     /product="terminal protein LMP2B"
                     /protein_id="CAD53383.1"
                     /db_xref="GI:23893578"
                     /db_xref="InterPro:IPR010881"
                     /db_xref="UniProt/TrEMBL:Q8AZK9"
                     /translation="MNPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLL
                     LAAVASSYAAAQRKLLTPVTVLTAVVTFFAICLTWRIEDPPFNSLLFALLAAAGGLQG
                     IYVLVMLVLLILAYRRRWRRLTVCGGIMFLACVLVLIVDAVLQLSPLLGAVTVVSMTL
                     LLLAFVLWLSSPGGLGTLGAALLTLAAALALLASLILGTLNLTTMFLLMLLWTLVVLL
                     ICSSCSSCPLSKILLARLFLYALALLLLASALIAGGSILQTNFKSLSSTEFIPNLFCM
                     LLLIVAGILFILAILTEWGSGNRTYGPVFMCLGGLLTMVAGAVWLTVMSNTLLSAWIL
                     TAGFLIFLIGFALFGVIRCCRYCCYYCLTLESEERPPTPYRNTV"
     exon            360. .458
                     /gene="LMP2"
                     /number=3
     exon            540. .788
                     /gene="LMP2"
                     /number=4
     exon            871. .951
                     /gene="LMP2"
                     /number=5
     exon            1026. .1196

  Query Match             98.8%;  Score 531.8;  DB 13;  Length 171823;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-92;
  Matches  533;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CATGATGGCACGCCGGCTGCCCAAGCCCACCCTCCAGGGGAGGCTGGAGGCGGATTTTCC 60
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49215 CGTTATGGCACGCCGGCTGCCCAAGCCCACCCTCCAGGGGAGGCTGGAGGCGGATTTTCC 49274

Qy         61 AGACAGTCCCCTGCTTCCTAAATTTCAAGAGCTGAACCAGAATAATCTCCCCAATGATGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49275 AGACAGTCCCCTGCTTCCTAAATTTCAAGAGCTGAACCAGAATAATCTCCCCAATGATGT 49334

Qy        121 TTTTCGGGAGGCTCAAAGAAGTTACCTGGTATTTCTGACATCCCAGTTCTGCTACGAAGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49335 TTTTCGGGAGGCTCAAAGAAGTTACCTGGTATTTCTGACATCCCAGTTCTGCTACGAAGA 49394

Qy        181 GTACGTGCAGAGGACTTTTGGGGTGCCTCGGCGCCAACGCGCCATAGACAAGAGGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49395 GTACGTGCAGAGGACTTTTGGGGTGCCTCGGCGCCAACGCGCCATAGACAAGAGGCAGAG 49454

Qy        241 AGCCAGTGTGGCTGGGGCTGGTGCTCATGCACACCTTGGCGGGTCATCCGCCACCCCCGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49455 AGCCAGTGTGGCTGGGGCTGGTGCTCATGCACACCTTGGCGGGTCATCCGCCACCCCCGT 49514

Qy        301 CCAGCAGGCTCAGGCCGCCGCATCCGCTGGGACCGGGGCCTTGGCATCATCAGCGCCGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49515 CCAGCAGGCTCAGGCCGCCGCATCCGCTGGGACCGGGGCCTTGGCATCATCAGCGCCGTC 49574

Qy        361 CACGGCCGTAGCCCAGTCCGCGACCCCCTCTGTTTCTTCATCTATTAGCAGCCTCCGGGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49575 CACGGCCGTAGCCCAGTCCGCGACCCCCTCTGTTTCTTCATCTATTAGCAGCCTCCGGGC 49634

Qy        421 CGCGACTTCGGGGGCGACTGCCGCCGCCTCCGCCGCCGCAGCCGTCGATACCGGGTCAGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49635 CGCGACTTCGGGGGCGACTGCCGCCGCCTCCGCCGCCGCAGCCGTCGATACCGGGTCAGG 49694
```

```
Qy        481 TGGCGGGGGACAACCCCACGACACCGCCCCACGCGGGGCACGTAAGAAACAGTAG 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49695 TGGCGGGGGACAACCCCACGACACCGCCCCACGCGGGGCACGTAAGAAACAGTAG 49749


RESULT 8
EBV
LOCUS       EBV                    172281 bp    DNA     circular VRL 18-APR-2005
DEFINITION  Epstein-Barr virus (EBV) genome, strain B95-8.
ACCESSION   V01555 J02070 K01729 K01730 V01554 X00498 X00499 X00784
VERSION     V01555.1  GI:59074
KEYWORDS    DNA polymerase; EBNA; genome; ribonucleotide reductase; tandem
            repeat; terminal repeat.
SOURCE      Human herpesvirus 4 (Epstein-Barr virus)
  ORGANISM  Human herpesvirus 4
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Gammaherpesvirinae; Lymphocryptovirus.
REFERENCE   1
  AUTHORS   Arrand,J.R., Rymo,L., Walsh,J.E., Bjorck,E., Lindahl,T. and
            Griffin,B.E.
  TITLE     Molecular cloning of the complete Epstein-Barr virus genome as a
            set of overlapping restriction endonuclease fragments
  JOURNAL   Nucleic Acids Res. 9 (13), 2999-3014 (1981)
   PUBMED   6269068
REFERENCE   2
  AUTHORS   Kozak,M.
  TITLE     Possible role of flanking nucleotides in recognition of the AUG
            initiator codon by eukaryotic ribosomes
  JOURNAL   Nucleic Acids Res. 9 (20), 5233-5252 (1981)
   PUBMED   7301588
REFERENCE   3
  AUTHORS   Deininger,P.L., Bankier,A., Farrell,P., Baer,R. and Barrell,B.
  TITLE     Sequence analysis and in vitro transcription of portions of the
            Epstein-Barr virus genome
  JOURNAL   J. Cell. Biochem. 19 (3), 267-274 (1982)
   PUBMED   6296170
REFERENCE   4
  AUTHORS   Farrell,P.J., Bankier,A., Seguin,C., Deininger,P. and Barrell,B.G.
  TITLE     Latent and lytic cycle promoters of Epstein-Barr virus
  JOURNAL   EMBO J. 2 (8), 1331-1338 (1983)
   PUBMED   10872327
REFERENCE   5
  AUTHORS   Farrell,P.J., Deininger,P.L., Bankier,A. and Barrell,B.
  TITLE     Homologous upstream sequences near Epstein-Barr virus promoters
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 80 (6), 1565-1569 (1983)
   PUBMED   6300857
REFERENCE   6  (bases 142687 to 159853)
  AUTHORS   Bankier,A.T., Deininger,P.L., Farrell,P.J. and Barrell,B.G.
  TITLE     Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8
            Epstein-Barr virus
  JOURNAL   Mol. Biol. Med. 1 (1), 21-45 (1983)
   PUBMED   6092825
REFERENCE   7  (bases 112620 to 125316)
  AUTHORS   Seguin,C., Farrell,P.J. and Barrell,B.G.
  TITLE     DNA sequence and transcription of the BamHI fragment B region of
            B95-8 Epstein-Barr virus
  JOURNAL   Mol. Biol. Med. 1 (3), 369-392 (1983)
   PUBMED   6094953
REFERENCE   8  (bases 45644 to 52450)
  AUTHORS   Jeang,K.T. and Hayward,S.D.
```